- The mean and **standard deviation** of a binomial random variable X are

$$\mu_X = np \qquad \sigma_X = \sqrt{np(1-p)}$$

- The binomial distribution with n trials and probability p of success gives a good approximation to the count of successes in an SRS of size n from a large population containing proportion p of successes. This is true as long as the sample size n is no more than 10% of the population size N (the **10% condition**).
- The **Normal approximation** to the binomial distribution* says that if X is a count of successes having the binomial distribution with parameters n and p, then when n is large, X is approximately Normally distributed with mean np and standard deviation $\sqrt{np(1-p)}$. We will use this approximation when $np \geq 10$ and $n(1-p) \geq 10$ (the **Large Counts condition**).
- A **geometric setting** consists of repeated trials of the same chance process in which the probability p of success is the same on each trial, and the goal is to count the number of trials it takes to get one success. If Y = the number of trials required to obtain the first success, then Y is a **geometric random variable.** Its probability distribution is called a **geometric distribution.**
- If Y has the geometric distribution with probability of success p, the possible values of Y are the positive integers 1, 2, 3, ... The **geometric probability** that Y takes any value is

$$P(Y = k) = (1-p)^{k-1}p$$

- The **mean** (expected value) of a geometric random variable Y is $\mu_Y = 1/p$.

*This topic is not required for the AP® Statistics exam. Some teachers prefer to discuss this topic when presenting the sampling distribution of $\hat{p}$ (Chapter 7).

6.3 TECHNOLOGY CORNERS

TI-Nspire Instructions in Appendix B; HP Prime instructions on the book's Web site.

Section 6.3 Exercises

In Exercises 69 to 72, determine whether the given random variable has a binomial distribution. Justify your answer.

69. **Sowing seeds** Seed Depot advertises that its new flower seeds have an 85% chance of germinating (growing). Suppose that the company's claim is true. Judy gets a packet with 20 randomly selected new flower seeds from Seed Depot and plants them in her garden. Let X = the number of seeds that germinate.

pg 388

70. **Long or short?** Put the names of all the students in your class in a hat. Mix them up, and draw four names without looking. Let Y = the number whose last names have more than six letters.

71. **Lefties** Exactly 10% of the students in a school are left-handed. Select students at random from the school, one at a time, until you find one who is left-handed. Let V = the number of students chosen.

72. **Taking the train** According to New Jersey Transit, the 8:00 A.M. weekday train from Princeton to New York City has a 90% chance of arriving on time on a randomly selected day. Suppose this claim is true. Choose 6 days at random. Let W = the number of days on which the train arrives late.

73. **Binomial setting?** A binomial distribution will be approximately correct as a model for one of these two settings and not for the other. Explain why by briefly discussing both settings.

(a) When an opinion poll calls residential telephone numbers at random, only 20% of the calls reach a person. You watch the random digit dialing machine make 15 calls. X is the number that reach a person.

(b) When an opinion poll calls residential telephone numbers at random, only 20% of the calls reach a live person. You watch the random digit dialing machine make calls. Y is the number of calls needed to reach a live person.

74. **Binomial setting?** A binomial distribution will be approximately correct as a model for one of these two sports settings and not for the other. Explain why by briefly discussing both settings.

(a) A National Football League kicker has made 80% of his field goal attempts in the past. This season he attempts 20 field goals. The attempts differ widely in distance, angle, wind, and so on.

(b) A National Basketball Association player has made 80% of his free-throw attempts in the past. This season he takes 150 free throws. Basketball free throws are always attempted from 15 feet away with no interference from other players.

75. **Elk** Biologists estimate that a baby elk has a 44% chance of surviving to adulthood. Assume this estimate is correct. Suppose researchers choose 7 baby elk at random to monitor. Let X = the number who survive to adulthood. Use the binomial probability formula to find $P(X = 4)$. Interpret this result in context. pg 393

76. **Rhubarb** Suppose you purchase a bundle of 10 bare-root rhubarb plants. The sales clerk tells you that 5% of these plants will die before producing any rhubarb. Assume that the bundle is a random sample of plants and that the sales clerk's statement is accurate. Let Y = the number of plants that die before producing any rhubarb. Use the binomial probability formula to find $P(Y = 1)$. Interpret this result in context.

77. **Elk** Refer to Exercise 75. How surprising would it be for more than 4 elk in the sample to survive to adulthood? Calculate an appropriate probability to support your answer. pg 393

78. **Rhubarb** Refer to Exercise 76. Would you be surprised if 3 or more of the plants in the bundle die before producing any rhubarb? Calculate an appropriate probability to support your answer.

79. **Sowing seeds** Refer to Exercise 69.

(a) Find the probability that exactly 17 seeds germinate. Show your work. pg 396

(b) If only 12 seeds actually germinate, should Judy be suspicious that the company's claim is not true? Compute $P(X \leq 12)$ and use this result to support your answer.

80. **Taking the train** Refer to Exercise 72.

(a) Find the probability that the train arrives late on exactly 2 days. Show your work.

(b) Would you be surprised if the train arrived late on 2 or more days? Compute $P(W \geq 2)$ and use this result to support your answer.

81. **Random digit dialing** When an opinion poll calls a residential telephone number at random, there is only a 20% chance that the call reaches a live person. You watch the random digit dialing machine make 15 calls. Let X = the number of calls that reach a live person.

(a) Find and interpret μ_X.

(b) Find and interpret σ_X.

82. **Lie detectors** A federal report finds that lie detector tests given to truthful persons have probability about 0.2 of suggesting that the person is deceptive.[12] A company asks 12 job applicants about thefts from previous employers, using a lie detector to assess their truthfulness. Suppose that all 12 answer truthfully. Let X = the number of people who the lie detector says are being deceptive.

(a) Find and interpret μ_X.

(b) Find and interpret σ_X.

83. **Random digit dialing** Refer to Exercise 81. Let Y = the number of calls that *don't* reach a live person.

(a) Find the mean of Y. How is it related to the mean of X? Explain why this makes sense.

(b) Find the standard deviation of Y. How is it related to the standard deviation of X? Explain why this makes sense.

84. **Lie detectors** Refer to Exercise 82. Let Y = the number of people who the lie detector says are telling the truth.

(a) Find $P(Y \geq 10)$. How is this related to $P(X \leq 2)$? Explain.

(b) Calculate μ_Y and σ_Y. How do they compare with μ_X and σ_X? Explain why this makes sense.

85. **1 in 6 wins** As a special promotion for its 20-ounce bottles of soda, a soft drink company printed a message on the inside of each cap. Some of the caps said, "Please try again," while others said, "You're a winner!" The company advertised the promotion with the slogan "1 in 6 wins a prize." Suppose the company is telling the truth and that every 20-ounce bottle of soda it fills has a 1-in-6 chance of being a winner. Seven friends each buy one 20-ounce bottle of the soda at a local convenience store. Let X = the number who win a prize.

pg 399

(a) Explain why X is a binomial random variable.

(b) Find the mean and standard deviation of X. Interpret each value in context.

(c) The store clerk is surprised when three of the friends win a prize. Is this group of friends just lucky, or is the company's 1-in-6 claim inaccurate? Compute $P(X \geq 3)$ and use the result to justify your answer.

86. **Aircraft engines** Engineers define reliability as the probability that an item will perform its function under specific conditions for a specific period of time. A certain model of aircraft engine is designed so that each engine has probability 0.999 of performing properly for an hour of flight. Company engineers test an SRS of 350 engines of this model. Let X = the number that operate for an hour without failure.

(a) Explain why X is a binomial random variable.

(b) Find the mean and standard deviation of X. Interpret each value in context.

(c) Two engines failed the test. Are you convinced that this model of engine is less reliable than it's supposed to be? Compute $P(X \leq 348)$ and use the result to justify your answer.

87. **Airport security** The Transportation Security Administration (TSA) is responsible for airport safety. On some flights, TSA officers randomly select passengers for an extra security check before boarding. One such flight had 76 passengers—12 in first class and 64 in coach class. Some passengers were surprised when none of the 10 passengers chosen for screening were seated in first class. Can we use a binomial distribution to approximate this probability? Justify your answer.

pg 402

88. **Scrabble** In the game of Scrabble, each player begins by drawing 7 tiles from a bag containing 100 tiles. There are 42 vowels, 56 consonants, and 2 blank tiles in the bag. Cait chooses her 7 tiles and is surprised to discover that all of them are vowels. Can we use a binomial distribution to approximate this probability? Justify your answer.

89. **10% condition** To use a binomial distribution to approximate the count of successes in an SRS, why do we require that the sample size n be no more than 10% of the population size N?

90. ***Large Counts condition** To use a Normal distribution to approximate binomial probabilities, why do we require that both np and $n(1-p)$ be at least 10?

91. ***On the Web** What kinds of Web sites do males aged 18 to 34 visit most often? Half of male Internet users in this age group visit an auction site such as eBay at least once a month.[13] A study of Internet use interviews a random sample of 500 men aged 18 to 34. Let X = the number in the sample who visit an auction site at least once a month.

pg 403

(a) Show that X is approximately a binomial random variable.

(b) Check the conditions for using a Normal approximation in this setting.

(c) Use a Normal distribution to estimate the probability that at least 235 of the men in the sample visit an online auction site at least once a month.

92. ***Checking for survey errors** One way of checking the effect of undercoverage, nonresponse, and other sources of error in a sample survey is to compare the sample with known facts about the population. About 12% of American adults identify themselves as black. Suppose we take an SRS of 1500 American adults and let X be the number of blacks in the sample.

(a) Show that X is approximately a binomial random variable.

(b) Check the conditions for using a Normal approximation in this setting.

(c) Use a Normal distribution to estimate the probability that the sample will contain between 165 and 195 blacks.

93. **Using Benford's law** According to Benford's law (Exercise 5, page 359), the probability that the first digit of the amount on a randomly chosen invoice is a 1 or a 2 is 0.477. Suppose you examine an SRS of 90 invoices from a vendor and find 29 that have first digits 1 or 2. Do you suspect that the invoice

*This topic is not required for the AP® Statistics exam. Some teachers prefer to discuss this topic when presenting the sampling distribution of $\hat{p}$ (Chapter 7).

amounts are not genuine? Compute an appropriate probability to support your answer.

94. **A .300 hitter** In baseball, a 0.300 hitter gets a hit in 30% of times at bat. When a baseball player hits 0.300, fans tend to be impressed. Typical Major Leaguers bat about 500 times a season and hit about 0.260. A hitter's successive tries seem to be independent. Could a typical Major Leaguer hit 0.300 just by chance? Compute an appropriate probability to support your answer.

95. **Geometric or not?** Determine whether each of the following scenarios describes a geometric setting. If so, define an appropriate geometric random variable.

(a) A popular brand of cereal puts a card with 1 of 5 famous NASCAR drivers in each box. There is a 1/5 chance that any particular driver's card ends up in any box of cereal. Buy boxes of the cereal until you have all 5 drivers' cards.

(b) In a game of 4-Spot Keno, Lola picks 4 numbers from 1 to 80. The casino randomly selects 20 winning numbers from 1 to 80. Lola wins money if she picks 2 or more of the winning numbers. The probability that this happens is 0.259. Lola decides to keep playing games of 4-Spot Keno until she wins some money.

96. **Geometric or not?** Determine whether each of the following scenarios describes a geometric setting. If so, define an appropriate geometric random variable.

(a) Shuffle a standard deck of playing cards well. Then turn over one card at a time from the top of the deck until you get an ace.

(b) Lawrence likes to shoot a bow and arrow in his free time. On any shot, he has about a 10% chance of hitting the bull's-eye. As a challenge one day, Lawrence decides to keep shooting until he gets a bull's-eye.

97. **1-in-6 wins** Alan decides to use a different strategy for the 1-in-6 wins game of Exercise 85. He keeps buying one 20-ounce bottle of the soda at a time until he gets a winner.

pg 406

(a) Find the probability that he buys exactly 5 bottles. Show your work.

(b) Find the probability that he buys no more than 8 bottles. Show your work.

98. **Cranky mower** To start her old lawn mower, Rita has to pull a cord and hope for some luck. On any particular pull, the mower has a 20% chance of starting.

(a) Find the probability that it takes her exactly 3 pulls to start the mower. Show your work.

(b) Find the probability that it takes her more than 10 pulls to start the mower. Show your work.

99. **Using Benford's law** According to Benford's law (Exercise 5, page 359), the probability that the first digit of the amount of a randomly chosen invoice is an 8 or a 9 is 0.097. Suppose you examine randomly selected invoices from a vendor until you find one whose amount begins with an 8 or a 9.

(a) How many invoices do you expect to examine until you get one that begins with an 8 or 9? Justify your answer.

(b) In fact, you don't get an amount starting with an 8 or 9 until the 40th invoice. Do you suspect that the invoice amounts are not genuine? Compute an appropriate probability to support your answer.

100. **Roulette** Marti decides to keep placing a \$1 bet on number 15 in consecutive spins of a roulette wheel until she wins. On any spin, there's a 1-in-38 chance that the ball will land in the 15 slot.

(a) How many spins do you expect it to take until Marti wins? Justify your answer.

(b) Would you be surprised if Marti won in 3 or fewer spins? Compute an appropriate probability to support your answer.

Multiple choice: Select the best answer for Exercises 101 to 105.

101. Joe reads that 1 out of 4 eggs contains salmonella bacteria. So he never uses more than 3 eggs in cooking. If eggs do or don't contain salmonella independently of each other, the number of contaminated eggs when Joe uses 3 chosen at random has the following distribution:

(a) binomial; $n = 4$ and $p = 1/4$

(b) binomial; $n = 3$ and $p = 1/4$

(c) binomial; $n = 3$ and $p = 1/3$

(d) geometric; $p = 1/4$

(e) geometric; $p = 1/3$

Exercises 102 and 103 refer to the following setting. A fast-food restaurant runs a promotion in which certain food items come with game pieces. According to the restaurant, 1 in 4 game pieces is a winner.

102. If Jeff gets 4 game pieces, what is the probability that he wins exactly 1 prize?

(a) 0.25

(b) 1.00

(c) $\binom{4}{1}(0.25)^1(0.75)^3$

(d) $\binom{4}{1}(0.25)^3(0.75)^1$

(e) $(0.75)^3(0.25)^1$

103. If Jeff keeps playing until he wins a prize, what is the probability that he has to play the game exactly 5 times?

(a) $(0.25)^5$

(b) $(0.75)^4$

(c) $(0.75)^5$

(d) $(0.75)^4(0.25)$

(e) $\binom{5}{1}(0.75)^4(0.25)$

104. Each entry in a table of random digits like Table D has probability 0.1 of being a 0, and the digits are independent of one another. If many lines of 40 random digits are selected, the mean and standard deviation of the number of 0s will be approximately

(a) mean = 0.1, standard deviation = 0.05.

(b) mean = 0.1, standard deviation = 0.1.

(c) mean = 4, standard deviation = 0.05.

(d) mean = 4, standard deviation = 1.90.

(e) mean = 4, standard deviation = 3.60.

105. *In which of the following situations would it be appropriate to use a Normal distribution to approximate probabilities for a binomial distribution with the given values of n and p?

(a) $n = 10, p = 0.5$

(b) $n = 40, p = 0.88$

(c) $n = 100, p = 0.2$

(d) $n = 100, p = 0.99$

(e) $n = 1000, p = 0.003$

*This topic is not required for the AP® Statistics exam. Some teachers prefer to discuss this topic when presenting the sampling distribution of $\hat{p}$ (Chapter 7).

106. **Spoofing** (4.2) To collect information such as passwords, online criminals use "spoofing" to direct Internet users to fraudulent Web sites. In one study of Internet fraud, students were warned about spoofing and then asked to log in to their university account starting from the university's home page. In some cases, the login link led to the genuine dialog box. In others, the box looked genuine but in fact was linked to a different site that recorded the ID and password the student entered. The box that appeared for each student was determined at random. An alert student could detect the fraud by looking at the true Internet address displayed in the browser status bar, but most just entered their ID and password. Is this study an experiment? Why? What are the explanatory and response variables?

107. **Smoking and social class** (5.3) As the dangers of smoking have become more widely known, clear class differences in smoking have emerged. British government statistics classify adult men by occupation as "managerial and professional" (43% of the population), "intermediate" (34%), or "routine and manual" (23%). A survey finds that 20% of men in managerial and professional occupations smoke, 29% of the intermediate group smoke, and 38% in routine and manual occupations smoke.[14]

(a) Use a tree diagram to find the percent of all adult British men who smoke.

(b) Find the percent of male smokers who have routine and manual occupations.

FRAPPY! Free Response AP® Problem, Yay!

The following problem is modeled after actual AP® Statistics exam free response questions. Your task is to generate a complete, concise response in 15 minutes.

Directions: Show all your work. Indicate clearly the methods you use, because you will be scored on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

Buckley Farms produces homemade potato chips that it sells in bags labeled "16 ounces." The total weight of each bag follows an approximately Normal distribution with a mean of 16.15 ounces and a standard deviation of 0.12 ounces.

(c) Buckley Farms ships its chips in boxes that contain 6 bags. The empty boxes have a mean weight of 10 ounces and a standard deviation of 0.05 ounces. Calculate the mean and standard deviation of the total weight of a box containing 6 bags of chips.

(d) Buckley Farms decides to increase the mean weight of each bag of chips so that only 5% of the bags have weights that are less than 16 ounces. Assuming that the standard deviation remains 0.12 ounces, what mean weight should Buckley Farms use?

After you finish, you can view two example solutions on the book's